ON GRADIENT CALCULATION IN QUATERNION ADAPTIVE FILTERING

Cyrus Jahanchahi[†], Clive Cheong Took[†], and Danilo P. Mandic[†]

[†] Electrical and Electronic Engineering Department, Imperial College London, SW7 2AZ, UK
E-mails: {cyrus.jahanchahi, c.cheong-took, d.mandic}@imperial.ac.uk

ABSTRACT

A novel way to calculate the gradient of real functions of quaternion variables, typical cost functions in quaternion signal processing, is proposed. This is achieved by revisiting quaternion involutions and by simplifying the existing $\mathbb{HR}$ derivatives. This has allowed us to express the class of quaternion least mean square (QLMS) algorithms in a more compact form while keeping the same generic form of LMS. Simulations in the prediction setting support the approach.

Index Terms— Quaternion gradient, conjugate gradient, involutions, HR calculus, Quaternion LMS (QLMS)

1. INTRODUCTION

Recent developments in sensor technology and human centred computing have brought to light a new class of three- and four-dimensional vector-valued signals. These are most conveniently described by quaternions, benefiting from precise notions of rotation and orientation, enabled by their division algebra. The growing interest in quaternion signal processing has highlighted the need to develop fundamental tools so as to exploit its full potential [1]. Recent advances in this direction include those in spectrum analysis [2], algebraic matrix decompositions [3] [4] and augmented statistics [5] [6]. However, the calculus of quaternions, especially in the context of the derivatives of non-analytic functions, such as the standard instantaneous squared error $J = |e|^2$ has only recently been addressed [7].

Statistical signal processing often involves optimisation of real-valued cost functions of quaternion variables which are not analytic, and hence cannot be differentiated in the usual 'complex' or 'quaternion' sense. The $\mathbb{CR}$ (Wirtinger) calculus addresses this problem in the complex domain $\mathbb{C}$ [8] [9], whereas it was not until only recently that the $\mathbb{HR}$-calculus provided a unifying way to differentiate real-valued functions in $\mathbb{H}$ [7]. The Wirtinger and $\mathbb{HR}$-calculus thus offer an elegant way to differentiate directly in the complex or hypercomplex division algebras, and thus bypass the cumbersome task of differentiating separately with respect to the real and imaginary parts of the corresponding variables. Following our work on $\mathbb{HR}$-calculus [7], we here extend the 'differentiability' framework to cater specifically for functions described by quaternion involutions $q^\eta \; \forall \eta \; \in \; \{\imath, \jmath, \kappa\}$ (analogous to the complex conjugate operator). The widely linear (WL) quaternion model [5] [6]

$$y = \mathbf{u}^H\mathbf{x} + \mathbf{v}^H\mathbf{x}^\imath + \mathbf{g}^H\mathbf{x}^\jmath + \mathbf{h}^H\mathbf{x}^\kappa \tag{1}$$

has made it increasingly apparent that quaternion involutions play a prominent role in quaternion signal processing. This model has been used, e.g. in adaptive filters to account for the complete second order statistics of quaternion variables [5]. Our aim is to introduce a simplified yet accurate way to differentiate with respect to quaternion involutions, and to propose a new gradient definition termed the 'i-gradient'. It is shown that the i-gradient not only points in the same direction as the conjugate gradient $\frac{\partial f(\cdot)}{\partial q^*}$ (standard pseudogradient and $\mathbb{HR}^*$−derivative), but also produces the same generic solutions as those in the real and complex domains. We illustrate the use of the $\mathbb{HR}$ calculus and the i-gradient in the context of the quaternion least mean square (QLMS) algorithm [10]. This is supported by a rigorous comparative analysis of the adaptive filters corresponding to the different gradient definitions, and a set of simulations, demonstrating the advantage of the i-gradient over the standard conjugate gradient.

2. QUATERNIONS AND ITS INVOLUTIONS

For a quaternion $q = q_a + \imath q_b + \jmath q_c + \kappa q_d = Sq + Vq$, the scalar (real) part is denoted by $q_a = Sq = \Re(q)$, whereas the vector part (also called pure quaternion) $Vq = \Im(q) = \imath q_b + \jmath q_c + \kappa q_d$ comprises the three imaginary parts. The noncommutative quaternion product is given by

$$q_1 q_2 = Sq_1 Sq_2 - Vq_1 \cdot Vq_2 + Sq_2 Vq_1 + Sq_1 Vq_2 + Vq_1 \times Vq_2$$

where the symbol '$\cdot$' denotes the scalar product[1] and '$\times$' the vector product. The quaternion conjugate is given by $q^* = Sq - Vq$, and the norm by $\| q \| = \sqrt{<q,q>} = \sqrt{qq^*}$, and thus $q^{-1} = \frac{q^*}{qq^*} = \frac{q^*}{\|q\|^2} = \frac{q^*}{<q,q>}$, and $\imath^{-1} = -\imath, \jmath^{-1} = -\jmath, \kappa^{-1} = -\kappa$.

---

[1] The scalar product $q_1 \cdot q_2 = <q_1, q_2>, q_1, q_2 \in \mathbb{H}$ is defined as

$$q_1 \cdot q_2 = q_{1a}q_{2a} + q_{1b}q_{2b} + q_{1c}q_{2c} + q_{1d}q_{2d} = \Re(q_1 q_2^*) = \Re(q_1^* q_2).$$

2.1. Similarity Relations and Involutions

The similarity relation between quaternions p and q denoted by '$\sim$' implies the following relationship

$$p \sim q \quad \Leftrightarrow \quad p = \mu q \mu^{-1}, \quad \mu \neq 0 \qquad (2)$$

Similarity relations of particular interest are the quaternion involutions[2] (self-inverse mappings)

$$q^\imath = q_a + \imath q_b - \jmath q_c - \kappa q_d \qquad q^\jmath = q_a - \imath q_b + \jmath q_c - \kappa q_d$$
$$q^\kappa = q_a - \imath q_b - \jmath q_c + \kappa q_d \qquad (4)$$

Notice that the quaternion conjugate is also an involution, that is, $(q^*)^* = q$. Properties of quaternion involutions for quaternion product can be summarised as

$$\begin{aligned}
(p^\eta)^\eta &= p & (p^\eta)^* &= (p^*)^\eta \\
(pq)^\eta &= p^\eta q^\eta & (pq)^{\eta*} &= q^{\eta*} p^{\eta*} \\
(p^\delta)^\eta &= (p^\eta)^\delta = p^\alpha & \forall \eta &\neq \delta \neq \alpha \in \{\imath, \jmath, \kappa\}
\end{aligned} \qquad (5)$$

The relationship between the involutions and the conjugate of a quaternion variable is given by

$$q^* = \frac{1}{2}(q^\imath + q^\jmath + q^\kappa - q) \qquad (6)$$

$$q = \frac{1}{2}(q^{\imath*} + q^{\jmath*} + q^{\kappa*} - q) \qquad (7)$$

The identities in (6) - (7) are crucial in the analysis of quaternion gradients, as discussed below.

2.2. The $\mathbb{HR}$-Calculus and Quaternion Differentiability

The $\mathbb{HR}$-calculus provides an enabling tool for the differentiation of both analytic and non-analytic functions of quaternion variables under the same umbrella. It was developed in [7] to circumvent the stringent Cauchy-Riemann-Fueter (CRF) conditions, which are satisfied only by linear functions and constants. The $\mathbb{HR}$ and $\mathbb{HR}^*$-derivatives are given respectively as [7]

$$\frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q} = \frac{1}{4}\left[\frac{\partial f}{\partial q_a} - \imath\frac{\partial f}{\partial q_b} - \jmath\frac{\partial f}{\partial q_c} - \kappa\frac{\partial f}{\partial q_d}\right] \quad (8)$$

$$\frac{\partial f(q^*, q^{\imath*}, q^{\jmath*}, q^{\kappa*})}{\partial q^*} = \frac{1}{4}\left[\frac{\partial f}{\partial q_a} + \imath\frac{\partial f}{\partial q_b} + \jmath\frac{\partial f}{\partial q_c} + \kappa\frac{\partial f}{\partial q_d}\right] \quad (9)$$

For instance, to use the $\mathbb{HR}$-derivative, we first express the function $f(\cdot)$ in terms of the involutions $q, q^\imath, q^\jmath, q^\kappa$ using (6) and then differentiate with respect to q. Consider the function $f(\cdot) = q^*$, its $\mathbb{HR}$-derivative is $\frac{\partial f}{\partial q} = -1/2$, due to the relationship in (6). This derivative is not possible to make using the CRF conditions, yet functions of the type $J = ee^* = |e|^2$ are standard in stochastic gradient signal processing.

---

[2]Each component of q can be expressed in terms of the involutions as [7]

$$q_a = \frac{1}{4}[q + q^\imath + q^\jmath + q^\kappa] \qquad q_c = \frac{1}{4\jmath}[q - q^\imath + q^\jmath - q^\kappa]$$

$$q_b = \frac{1}{4\imath}[q + q^\imath - q^\jmath - q^\kappa] \qquad q_d = \frac{1}{4\kappa}[q - q^\imath - q^\jmath + q^\kappa] \quad (3)$$

2.3. The Proposed Quaternion Gradient

The gradient obtained from the $\mathbb{HR}$-derivatives [7] is rigorous and enables the derivation of stochastic gradient algorithms in the quaternion domain with respect to both q and its conjugate q^*. However, the relationships in (4) and (6) suggest that quaternion gradients with respect to the involutions should also be accounted for. We shall now consider an alternative gradient definition as follows. From (6) and (7), the partial derivative $\frac{\partial f}{\partial q}$ and its conjugate $(\frac{\partial f}{\partial q})^*$ can be written as

$$\frac{\partial f}{\partial q} = \frac{1}{2}\left(\left(\frac{\partial f}{\partial q}\right)^{\imath*} + \left(\frac{\partial f}{\partial q}\right)^{\jmath*} + \left(\frac{\partial f}{\partial q}\right)^{\kappa*} - \left(\frac{\partial f}{\partial q}\right)^*\right)$$

$$\left(\frac{\partial f}{\partial q}\right)^* = \frac{1}{2}\left(\left(\frac{\partial f}{\partial q}\right)^\imath + \left(\frac{\partial f}{\partial q}\right)^\jmath + \left(\frac{\partial f}{\partial q}\right)^\kappa - \frac{\partial f}{\partial q}\right)$$

Observe that for a real function $f(q) \in \mathbb{H}$ we have $\frac{\partial f}{\partial q} = \left(\frac{\partial f(q)}{\partial q^*}\right)^*$, and the above equations can be rewritten as

$$\frac{\partial f}{\partial q} = \frac{1}{2}\left(\frac{\partial f}{\partial q^{\imath*}} + \frac{\partial f}{\partial q^{\jmath*}} + \frac{\partial f}{\partial q^{\kappa*}} - \frac{\partial f}{\partial q^*}\right) \qquad (10)$$

$$\frac{\partial f}{\partial q^*} = \frac{1}{2}\left(\frac{\partial f}{\partial q^\imath} + \frac{\partial f}{\partial q^\jmath} + \frac{\partial f}{\partial q^\kappa} - \frac{\partial f}{\partial q}\right) \qquad (11)$$

It was shown in [7] that in optimization problems in $\mathbb{H}$, it is the conjugate gradient $\nabla_{q^*} f$ rather than the standard gradient $\nabla_q f$ that should be used, for it provides the maximum rate of change of $f(q)$. We therefore proceed by substituting (10) into (11) to give

$$\frac{\partial f}{\partial q^*} = \frac{1}{2}\left(\frac{\partial f}{\partial q^\imath} + \frac{\partial f}{\partial q^\jmath} + \frac{\partial f}{\partial q^\kappa}\right.$$
$$\left. - \frac{1}{2}\left(\frac{\partial f}{\partial q^{\imath*}} + \frac{\partial f}{\partial q^{\jmath*}} + \frac{\partial f}{\partial q^{\kappa*}} - \frac{\partial f}{\partial q^*}\right)\right) \quad (12)$$

yielding after factorisation

$$\frac{3}{4}\frac{\partial f}{\partial q*} = \frac{1}{2}\left(\frac{1}{2}\left(\frac{\partial f}{\partial q^\imath} + \frac{\partial f}{\partial q^\jmath} + \frac{\partial f}{\partial q^\kappa}\right) + \frac{1}{2}\left(\frac{\partial f}{\partial q^\imath} - \frac{\partial f}{\partial q^{\imath*}}\right)\right.$$
$$\left. + \frac{1}{2}\left(\frac{\partial f}{\partial q^\jmath} - \frac{\partial f}{\partial q^{\jmath*}}\right) + \frac{1}{2}\left(\frac{\partial f}{\partial q^\kappa} - \frac{\partial f}{\partial q^{\kappa*}}\right)\right) \quad (13)$$

Upon using the following identities

- $\frac{1}{2}(q - q^*) = \Im(q)$ where the operator $\Im(\cdot)$ stands for the imaginary (vector) part of q

- $\frac{\partial f(q)}{\partial q^*} = \frac{\partial f^*(q)}{\partial q^*} = \left(\frac{\partial f(q)}{\partial q}\right)^*$ valid for a real function of quaternion variables $f(q)$

the partial derivative $\frac{\partial f}{\partial q^*}$ in (13) becomes

$$\frac{\partial f}{\partial q^*} = \frac{1}{3}\left(\left(\frac{\partial f}{\partial q^\imath} + \frac{\partial f}{\partial q^\jmath} + \frac{\partial f}{\partial q^\kappa}\right)\right.$$
$$\left. + 2\left(\Im\left(\frac{\partial f}{\partial q^\imath}\right) + \Im\left(\frac{\partial f}{\partial q^\jmath}\right) + \Im\left(\frac{\partial f}{\partial q^\kappa}\right)\right)\right) \quad (14)$$

that is, as desired, as a function of the involutions $\frac{\partial f}{\partial q^i}$, $\frac{\partial f}{\partial q^j}$ and $\frac{\partial f}{\partial q^\kappa}$. This way, the involution-wise partial derivatives $\frac{\partial f}{\partial q^i}$, $\frac{\partial f}{\partial q^j}$ and $\frac{\partial f}{\partial q^\kappa}$ are all components of the gradient $\frac{\partial f}{\partial q^*}$, whose real and imaginary parts are

$$\Re\left[\frac{\partial f}{\partial q^*}\right] = \frac{1}{3}\Re\left[\frac{\partial f}{\partial q^i} + \frac{\partial f}{\partial q^j} + \frac{\partial f}{\partial q^\kappa}\right] \quad (15)$$

$$\Im\left[\frac{\partial f}{\partial q^*}\right] = \Im\left[\frac{\partial f}{\partial q^i} + \frac{\partial f}{\partial q^j} + \frac{\partial f}{\partial q^\kappa}\right] \quad (16)$$

Observe that the direction of the vector part of the gradient $\frac{\partial f}{\partial q^*}$ is equivalent to $\frac{\partial f}{\partial q^i} + \frac{\partial f}{\partial q^j} + \frac{\partial f}{\partial q^\kappa}$, the only difference being the magnitude of the real part. This motivates us to adopt a new definition of quaternion gradient, termed the involution-or i-gradient, given by two equivalent expressions

$$\nabla_{w^\eta} J = \frac{\partial J}{\partial q^i} + \frac{\partial J}{\partial q^j} + \frac{\partial J}{\partial q^\kappa} \quad (17)$$

$$\nabla_{w^\eta} J = \frac{1}{4}\left[3\frac{\partial}{\partial q_a} + i\frac{\partial f}{\partial q_b} + j\frac{\partial f}{\partial q_c} + \kappa\frac{\partial f}{\partial q_d}\right] = \frac{\partial f}{\partial q^*} + \frac{1}{2}\frac{\partial f}{\partial q_a} \quad (18)$$

Remark#1: As compared with $\mathbb{HR}$ calculus, The i-gradient in (18) includes an additional term $0.5\partial f/\partial q_a$, which increases the steepness of the descent when minimising the real-valued function $f(\cdot)$ with respect to the real component of the quaternion variable q_a.

Remark#2: This suggests a potential faster convergence of the i-gradient, as compared to the standard conjugate gradient derived from the $\mathbb{HR}$ calculus.

3. iQUATERNION LEAST MEAN SQUARE ALGORITHM

We now demonstrate the usefulness of the i-gradient in adaptive filtering applications. Consider the gradient of the cost function $J(k) = \frac{1}{2}e(k)e^*(k)$ in the form

$$\nabla_{\mathbf{w}^\eta} J(k) = \frac{1}{2}\sum_{\eta=\{i,j,\kappa\}} e(k)\frac{\partial e^*(k)}{\partial \mathbf{w}^\eta(k)} + \frac{\partial e(k)}{\partial \mathbf{w}^\eta(k)}e^*(k) \quad (19)$$

Using the rules of $\mathbb{HR}$-calculus, we have [7]

$$\frac{\partial e^*(k)}{\partial \mathbf{w}^\eta} = \mathbf{0} \quad \forall \eta \in \{i,j,\kappa\} \quad (20)$$

To obtain $\frac{\partial e(k)}{\partial \mathbf{w}^i}, \frac{\partial e(k)}{\partial \mathbf{w}^j}$ and $\frac{\partial e(k)}{\partial \mathbf{w}^\kappa}$, substitute $\mathbf{w}^*$ into the expression for $e(k)$, to give

$$e(k) = d(k) - \frac{1}{2}\left(\mathbf{w}^i(k) + \mathbf{w}^j(k) + \mathbf{w}^\kappa(k) - \mathbf{w}(k)\right)\mathbf{x}(k)$$

The gradients $\frac{\partial e(k)}{\partial \mathbf{w}^i}, \frac{\partial e(k)}{\partial \mathbf{w}^j}$ and $\frac{\partial e(k)}{\partial \mathbf{w}^\kappa}$ can now be calculated by direct differentiation, yielding

$$\frac{\partial e(k)}{\partial \mathbf{w}^\eta} = -\frac{1}{2}\mathbf{x}(k) \quad \forall \eta \in \{i,j,\kappa\} \quad (21)$$

Substituting (20) and (21) into (19) and using the result in the steepest descent $\mathbf{w}(k+1) = \mathbf{w}(k) - \mu\nabla_{\mathbf{w}^\eta} J(k)$ gives the following weight update

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{3}{4}\mu\mathbf{x}(k)e^*(k) \quad (22)$$

This algorithm is termed the iQLMS (the i-gradient based QLMS). Observe that it has the same generic form as the LMS and complex LMS [9], as the term $\frac{3}{4}$ can be absorbed into the learning rate, and that the iQLMS requires half the computations of the QLMS, compare the updates in (24) and (26). In the same spirit, the WL-iQLMS based on the widely linear version of QLMS can be derived as [10]

$$\mathbf{w}^a(k+1) = \mathbf{w}^a(k) + \frac{3}{4}\mu\mathbf{x}^a(k)e^*(k) \quad (23)$$

where $\mathbf{x}^a = [\mathbf{x}^T \ \mathbf{x}^{iT} \ \mathbf{x}^{jT} \ \mathbf{x}^{kT}]^T$ is the augmented input vector and $\mathbf{w}^a = [\mathbf{u}^T \ \mathbf{v}^T \ \mathbf{g}^T \ \mathbf{h}^T]^T$ in the WL model (1).

3.1. Comparative Analysis of Quaternion Adaptive Filters

The original expression for the QLMS weight update $\Delta\mathbf{w} = \mathbf{w}(k+1) - \mathbf{w}(k)$, is given by [10]

$$\Delta\mathbf{w} = \mu\left(\frac{1}{2}e(k)\mathbf{x}^*(k) - \frac{1}{4}\mathbf{x}^*(k)e^*(k)\right) \quad (24)$$

while the expression for the QLMS update derived using the $\mathbb{HR}$-calculus based gradient $\nabla_{\mathbf{w}^*} J$, is

$$\Delta\mathbf{w} = \mu\left(\frac{1}{2}e(k)\mathbf{x}^*(k) - \frac{1}{4}\mathbf{x}(k)e^*(k)\right) \quad (25)$$

and for the iQLMS algorithm, derived using the involution derivatives, the update is given by

$$\Delta\mathbf{w} = \frac{3}{4}\mu e(k)\mathbf{x}^*(k) \quad (26)$$

It can be shown that the updates in (24) – (26) can be expressed in the same generic form as

$$\Re[\Delta\mathbf{w}] = \alpha\,\mu\Re\left[e(k)\Re[\mathbf{x}(k)]\right] - \beta\,\mu\Re\left[e(k)\Im[\mathbf{x}(k)]\right]$$

$$\Im[\Delta\mathbf{w}] = \gamma\,\mu\Im\left[e(k)\Re[\mathbf{x}(k)]\right] - \delta\,\mu\Im\left[e(k)\Im[\mathbf{x}(k)]\right] \quad (27)$$

The weight updates for the two QLMS and iQLMS only differ in the weighting factors as shown in Table 1.

Remark#3: The geometries of updates for all the three algorithms in (24) – (26) are topologically similar, as the weight updates follow the same principle.

Remark#4: The same information is used in the weight updates of all the three algorithms, indicating that they should exhibit similar steady state performances.

Remark#5: The difference in the coefficients α, β, γ, and δ in (27) suggests different rates of convergence for (24) – (26), with the iQLMS having the fastest convergence rate.
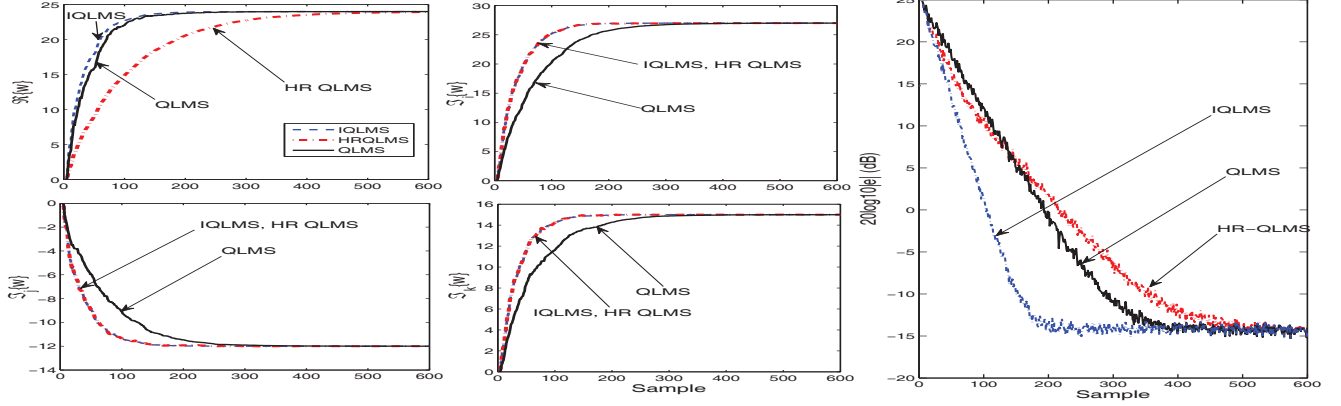
Fig. 1. Left, Center: Evolutions of adaptive weights of QLMS, HR-QLMS and iQLMS for one trial on the prediction of the MA(4) signal in (28). Right: Learning curves of the QLMS, HR-QLMS and iQLMS averaged over 50 trials of the signal in (28).

Table 1. Comparative analysis of quaternion adaptive filters

Algorithms in their generic form (27)	α	β	γ	δ
QLMS in (24)	1/4	3/4	3/4	1/4
$\mathbb{HR}$-QLMS in (25)	1/4	1/4	3/4	3/4
iQLMS in (26)	3/4	3/4	3/4	3/4

4. SIMULATIONS

The convergence properties of the QLMS, $\mathbb{HR}$-QLMS, and iQLMS were assessed on 50 averaged trials shown in the right plot of Fig. 1. In this set of simulations, the adaptive filters were set in a prediction setting and a moving average MA(4) signal employed is given by

$$y(k) = w_0 x(k) + \ldots + w_4 x(k-4) + n(k) \qquad (28)$$

where $w_0, \ldots, w_4$ are quaternion-valued weights initialised randomly from a quadrivariate Gaussian distribution and $n(k)$ is quadruply white Gaussian circular noise. Observe that the iQLMS and QLMS algorithms offer faster convergence than the $\mathbb{HR}$-QLMS, conforming with Remark 4. Remark 2 and 3 explain the reasons why all the three QLMS algorithms exhibited the same performance in the steady state. The interpretation of the four left hand plots in Fig. 1 is that the better convergence behaviour of iQLMS over $\mathbb{HR}$-QLMS for the real part $\Re\{\mathbf{w}\}$ of the weight can be attributed to the additional gradient component $\frac{\partial f}{\partial \mathbf{q}_a}$ in (18).

5. CONCLUSIONS

We have extended the $\mathbb{HR}$-calculus to enable gradient computations with respect to the quaternion involutions. Analysis has shown that the i-gradient points in the same direction of the conjugate gradient and includes an additional gradient component with respect to the real part of a quaternion

variable. Moreover, in the context of the QLMS algorithm, the use of the i-gradient has simplified its update, resulting in half the computational complexity. Comparative analysis has shown that QLMS, $\mathbb{HR}$-QLMS and iQLMS are essentially the same (and thus have the same steady state performance), however the iQLMS has the same generic form as LMS and CLMS.

6. REFERENCES

[1] J. Via, N. Le Bihan, D. Mandic, and S. Sangwine, "Quaternion signal processing: Theory and applications," *Tutorial in EUSIPCO*, 2011.

[2] S. Said, N. L. Bihan, and S. J. Sangwine, "Fast complexified quaternion Fourier transform," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1522–1531, 2008.

[3] N. L. Bihan and J. Mars, "Singular value decomposition of quaternion matrices: A new tool for vector-sensor signal processing," *Signal Processing*, vol. 84, no. 7, pp. 1177 – 1199, 2004.

[4] C. Cheong Took, D. P. Mandic, and F. Zhang, "On the Unitary Diagonalisation of a Special Class of Quaternion Matrices," *Applied Mathematics Letters*, vol. 24, no. 11, pp. 1806 – 1809, 2011.

[5] C. Cheong Took and D. P. Mandic, "Augmented Second-Order Statistics of Quaternion Signals," *Signal Processing*, vol. 91, no. 2, pp. 214 – 224, 2011.

[6] J. Vía, D. Ramírez, and I. Santamaría, "Properness and Widely Linear Processing of Quaternion Random Vectors," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3502–3515, 2010.

[7] D. P. Mandic, C. Jahanchahi, and C. Cheong Took, "A quaternion gradient operator and its applications," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 47 –50, 2011.

[8] K. Kreutz-Delgado, "The complex gradient operator and the $\mathbb{CR}$ calculus," *Lecture Supplement ECE275A*, pp. 1–76, 2006.

[9] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, John Wiley and Sons Ltd, 2009.

[10] C. Cheong Took and D. P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1316–1327, 2009.